Research Report

Prevalence and Ascertainment of Dementia Cases in the Malmö Diet and Cancer Study

Katarina Nägga[a,b,*], Vanja Bränsvik[a], Erik Stomrud[a,c], Olle Melander[d], Peter M. Nilsson[d],
Anna-Märta Gustavsson[a,c] and Oskar Hansson[a,c]
[a]*Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University, Sweden*
[b]*Department of Acute Internal Medicine and Geriatrics, Linköping University, Linköping, Sweden*
[c]*Memory Clinic, Skåne University Hospital, Malmö, Sweden*
[d]*Department of Clinical Sciences Malmö, Lund University, Sweden*

Abstract.
Background: Register diagnoses, both hospital-based and from open clinic care, are often used in research studies in Sweden. The validity of such diagnoses has been debated and a validation assessment can improve the diagnostic accuracy for use in research studies.
Objective: The aim of this study was to investigate the quality of register-derived dementia diagnoses in the Malmö Diet and Cancer population study (MDCS) and to validate these diagnoses using systematic criteria.
Methods: MDCS is a population-based prospective study comprising 30,446 participants. Register diagnoses of dementia for the MDCS population were derived from the Swedish National Patient Register (NPR) and validated through re-evaluation of available medical records by physicians.
Results: In the MDCS cohort, 2,206 participants were diagnosed with dementia according to the NPR during a mean follow-up of 18.1 years. The general dementia diagnosis was valid in 96% of the cases, but 40% of the specific dementia diagnoses were changed during the process of reevaluation. The diagnostic validity varied between 25.2% and 82.9% for the different diagnoses. The results from the validity assessment per diagnostic category revealed that the validity of the NPR diagnoses was higher for the more specific diagnoses and lower for unspecified dementia. The major diagnostic shift during the re-evaluation was from unspecified dementia to more specific diagnoses.
Conclusion: Validation of dementia diagnoses using medical records results in more precise diagnoses. Dementia diagnoses derived from registers should be validated in order to study associations between influential factors and different dementia diagnoses.

Keywords: Alzheimer's disease, dementia, register, risk factors, validation study

INTRODUCTION

Modern research has focused on intervention on risk factors for the prevention or at least postponement of dementia onset [1, 2]. It has been indicated that a delay of dementia diagnosis through minimization of potential influential factors could result in a reduction of dementia prevalence in the future [3–5]. Active interventions focusing on lifestyle factors have shown to have a protective effect on overall cognition in at-risk elderly [6].

There is a large number of studies indicating that vascular risk factors such as hypertension and dyslipidemia are associated with dementia [7] and also

*Correspondence to: Katarina Nägga, MD, PhD, Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden. Tel.: +46 0 725 29 80 24; E-mail: katarina.nagga@med.lu.se.

Alzheimer's disease [8]. However, different studies show somewhat inhomogeneous results due to differences in methodology. One relevant methodological aspect is the observation time, which is a limiting factor in clinical studies as the follow up time ideally needs to be several years in studies of dementia. This is due to an asymptomatic preclinical state and the slow and successive progression of the disorder. Some studies have used a prospective approach with follow-up of the study cohort for identification of cases who developed dementia [9]. An alternative to longitudinal studies with long observation times is to use register-based dementia diagnoses to follow-up cases in cohort studies where different risk factors have been assessed previously, and thus observe the number of incident dementia diagnoses over time. One challenge with this method is that dementia still tends to be underdiagnosed, and all cases may not have undergone a systematic and thorough dementia clinical work-up before diagnosis, which can result in an underestimation of true dementia prevalence and incorrect diagnoses. Hence, register diagnoses have been considered less well characterized than diagnoses derived from a research protocol, and therefore their validity for scientific studies has been debated.

In previous epidemiological studies of different risk factors for dementia diagnoses, the results vary, likely as a result of the use of different diagnostic criteria as well as differences in the diagnostic work-up. To be able to study differences between dementia subtypes, a correct diagnostic classification is therefore crucial.

In Sweden, there is a long tradition of the use of diagnostic national registers both in hospital-based care but also in open policlinic care. Register diagnoses are often used in research studies [10].

In the present study, we aimed to investigate the quality of the register-derived diagnoses by examining coherence to guidelines of diagnostic work-up, e.g., brain imaging, cognitive tests, and informant interview. Further we aimed to describe and thereby better define the dementia cohort within the large Malmö Diet and Cancer population study (MDCS). For this purpose, we performed a thorough procedure to validate the register diagnoses derived from the Swedish National Patient Register (NPR). Thus, trained physicians re-evaluated the available medical records to optimize diagnostic accuracy.

As previously described, different vascular risk factors have been associated with dementia diagnoses [7, 8]. A secondary aim of this study was therefore to describe baseline data of traditional risk markers

and protective factors for dementia in the established dementia cohort, to ensure the representativity of the cohort for future, more exploratory studies.

METHODS

Malmö Diet and Cancer study

MDCS is a population-based prospective study comprising 30,446 participants from the south of Sweden [11]. The baseline 1991–1996 examination encompassed a questionnaire on health status and different lifestyle factors from included individuals between the age of 44 and 74 years. A clinical investigation with measurement of blood pressure, height and weight, and blood sampling was also performed on the study participants. Previous publications have described the protocol and cohort characteristics in more detail [11–13]. Risk factor data in this study origins from the MDCS baseline assessments. Data on education, smoking status, and medication was based on self-reported information using a questionnaire. Hypertension was defined as a systolic blood pressure ≥ 140 mmHg and/or a diastolic blood pressure ≥ 90 mmHg at the baseline visit. Blood samples were collected to determine apolipoprotein E $\varepsilon 4$ allele status and height and weight was measured to calculate body mass index (BMI).

National patient register

All registered dementia diagnoses in the NPR through 31 December 2014 were obtained. The NPR started in the 1960s and since 1987 it covers 99% of all in-patient medical diagnoses automatically transferred to the register. Also, since 2001 it includes hospital-based (policlinic) out-patient diagnoses [14]. However, it does not cover diagnoses in the primary health care unless the diagnosis is registered during hospitalization or at a policlinic hospital visit. All diagnoses are routinely registered according to the International Classification of Diseases.

The following dementia diagnoses were obtained from the NPR for the present study; Alzheimer's disease dementia (AD) (F00, G30, 331A/331.0), vascular dementia (VaD) (F01, 290E/290.4), Parkinson's disease dementia (PDD) (F023), dementia with Lewy bodies (DLB) (F028, G318A), frontotemporal dementia (FTD) (F020, G310, 331B/331.1), or unspecified dementia (F03, 290, 294B/294.1, 331 C/331.2). The first register diagnosis per patient was subjected to diagnostic review described below.

A total of 2,206 dementia diagnoses were identified in the NPR.

Validation procedure

Next, the validity of the NPR dementia diagnoses was assessed using data from medical records (electronic charts) by trained physicians at the Memory Clinic, Skåne University Hospital in Malmö. Only existing dementia diagnoses extracted from the NPR were validated. The assessment of the diagnoses comprised key data defined as: 1) presentation of symptoms of cognitive impairment; 2) cognitive test results; and 3) computed tomography or magnetic resonance imaging of the brain. In the case of available cerebrospinal fluid (CSF) analyses, this was also included in the assessment but was not regarded as key data.

For the assessment of validity of each diagnosis, we categorized the amount of collected information from the medical records into five qualitative categories; 0 = almost non-existent (only information about diagnosis); 1 = very sparse (two of the key data missing); 2 = sparse (one of the key data missing), 3 = adequate (all key data present), and 4 = good (assessed at a tertiary clinic or equal level).

The diagnoses were then grouped into four different categories characterizing the validity of the diagnoses after the quality assessment of the data. These groups were; 0 = uncertain diagnosis (available data indicated different diagnoses to the same extent); 1 = possible diagnosis (some data indicated a specific diagnosis but important data was missing or the data did not fully fit into the same diagnostic criteria); 2 = probable diagnosis (a vast majority of the data indicated a specific diagnosis but some important data was missing or did not fully fit into the criteria); and 3 = definite diagnosis (all of the key data indicated the same specific diagnosis).

The criteria of *The Diagnostic and Statistical Manual of Mental Disorders*, fifth edition, were applied for the validated diagnoses [15]. The mixed dementia diagnosis was used when both AD pathology and cerebrovascular disease were assumed to significantly contribute to the clinical description. Unclear cases were subject to discussion with at least one very experienced senior dementia specialist resulting in a consensus diagnosis. Compared to the information available to the clinicians who initially gave the dementia diagnosis, the validation procedure allowed access to data covering a longer time span which could lead to a revised diagnosis based on symptom progression, repeated cognitive testing, or later added diagnostic data, such as brain imaging or CSF analyses. In several cases, unspecified dementia diagnosis from primary care could be revised by experienced dementia specialists involved in the study.

Risk and protective factors

Age [7, 16], sex, education, and presence of an apolipoprotein E ε4 allele are all known factors to be associated with dementia and have been used in previous studies [7]. Other variables included in the MDCS baseline data were body mass index, systolic blood pressure, smoking status, prevalent coronary event, prevalent stroke, prevalent diabetes, antihypertensive medication, lipid-lowering medication, and the ratio between serum-apolipoprotein B and serum-apolipoprotein A1 (S-ApoB/ApoA ratio), which was used as a measure of cholesterol levels. Hypertension was defined as a systolic blood pressure ≥ 140 mmHg and/or a diastolic blood pressure ≥ 90 mmHg regardless of antihypertensive medication. This definition was based on the definition by the World Health Organization [17], but in this study, only one measurement of blood pressure at baseline was used.

Statistics

Descriptive data from the validation process are presented as frequency with numbers and percentages or mean values and standard deviations. Group comparisons were performed using SPSS version 28.0. Group differences in numeric variables were estimated with Mann-Whitney U test and categorical variables with Pearson's χ^2 test. Two-tailed *p*-values < 0.05 were noted as statistically significant.

Ethical aspects

Before inclusion in the MDCS, all participants signed an informed consent. When dementia diagnoses were later collected from the NPR and merged to the MDCS for the validation process, advertisements were published in the local newspapers, enabling individuals to opt-out. The present study has been subject to ethical evaluation and was approved by the regional ethical review board in Lund (2011/83 and 2013/489).

Table 1
Results from the validation process of 2,206 dementia diagnoses from the National
Patient Register (NPR)

	N	%
Change of original NPR diagnosis	883	40.0
Availability of key data:		
1. Symptom presentation	1,904	86.3
2. Cognitive test results	1,894	85.9
3. Neuroradiology	2,080	94.3
Consulting specialist in cognitive disorders	1,697	76.9
Cerebrospinal fluid analyses	839	38.0
Quality of accessible data from medical records:		
0 = almost non-existent	96	4.4
1 = very sparse	145	6.6
2 = sparse	274	12.4
3 = adequate	314	14.2
4 = good (assessed at a tertiary clinic or equal level)	1,377	62.4
Validity of the diagnosis:		
0 = uncertain diagnosis	98	4.4
1 = possible diagnosis	323	14.6
2 = probable diagnosis	778	35.3
3 = definite diagnosis	1,007	45.6

RESULTS

Study population

The MDCS cohort comprise 30 446 participants and the mean (standard deviation (SD)) age at baseline was 58 (7.6) years. In all, 60.2% were women and 11.7% were born outside of Sweden.

Validation of diagnoses

In the MDCS cohort, 2,206 individuals had a registered dementia diagnosis according to NPR during a mean (SD) follow-up of 18.1 (5.1) years. Table 1 present results of the validation assessment, specifying diagnostic data availability, diagnostic quality, and diagnose validity category. After the validation assessment, 2,120 dementia diagnoses remained. Hence, 86 cases with dementia according to NPR were found not to meet the dementia criteria. Of these, there were 9 cases with no dementia, 17 with delirium, 16 with psychiatric diagnoses, 7 with other neurodegenerative disorders, 9 with cerebrovascular incidents, and 28 with mild cognitive impairment.

In total, 1,697 (76.9%) of the individuals with a NPR dementia diagnosis had been diagnosed by a specialist in cognitive disorders, either at a tertiary unit (Memory Clinic) or through a consultant assessment during hospital care. Of all patients, 1,904 (86.3%) had data on symptom presentation, 1,894 (85.9%) on cognitive test results, and 2,080 (94.3%) had undergone neuroimaging (either computed tomography or magnetic resonance imaging, or both). CSF analyses were performed in 839 (38.0%) of the participants (Table 1).

The 2,120 validated dementia diagnoses were distributed as follows in incidence numbers for various diagnoses; 621 (29.3%) AD, 595 (28.1%) mixed dementia (AD dementia in combination with cerebrovascular pathology; Mixed), 531 (25.0%) VaD, 96 (4.5%) DLB, 49 (2.3%) PDD, 26 (1.2%) FTD, 140 (6.6%) unspecified dementia (lack of information for categorization of origin), 15 (0.7%) normal pressure hydrocephalus, 9 (0.4%) alcohol-related dementia, 22 (1.0%) dementia with other mixed pathologies, 6 (0.3%) other neurodegenerative disorders, and 10 (0.5%) cases with atypical parkinsonism. The unspecified dementia diagnosis was applied in cases where a more specific dementia diagnosis could not be defined. The distribution of dementia diagnoses before and after the validation procedure is presented in Table 2, showing that the major change from the NPR diagnosis to the validated diagnosis was seen for the category unspecified dementia. The proportion of unchanged diagnoses after the validation process was as follows per diagnostic category: AD (63.5%), Mixed (82.9%), VaD (72.2%), DLB (63.0%), PDD (79.6%), FTD (65.4%), and unspecified dementia (25.2%).

The results from the validity assessment per diagnostic category revealed that the validity was higher for the more specific diagnoses, but lower for unspecified dementia (Table 3). For the more uncommon diagnoses such as DLB and FTD, there was a high

Table 2
Distribution of dementia diagnoses after validation of 2,206 register diagnoses from the National Patient Register (NPR)

	NPR diagnosis		Validated diagnosis after validation		Remaining NPR diagnosis after validation	
	n	(%)	*n*	(%)	*n*	(%)
Alzheimer's disease dementia	792	35.9	621	28.2	503	63.5
Mixed dementia	258	11.7	595	27.0	214	82.9
Vascular dementia	540	24.5	531	24.1	390	72.2
Dementia with Lewy Bodies	73	3.3	96	4.4	47	63.0
Parkinson disease with dementia	49	2.2	49	2.2	39	79.6
Frontotemporal dementia	26	1.2	26	1.2	17	65.4
Unspecified dementia	445	20.2	140	6.3	112	25.2
No dementia			86	3.9		
Normal Pressure Hydrocephalus			15	0.7		
Alcohol related dementia			9	0.4		
Mixed pathologies other			22	1.0		
Other neurodegenerative disorders	23	1.0	6	0.3		
Atypical parkinsonism			10	0.5		
Total	2206	100.0	2206	100.0		

"Remaining NPR diagnosis after validation" indicating the number of original diagnoses from the NPR which remained after validation.

quality of the information from the medical records which was also reflected in a high validity of the diagnoses. For unspecified dementia there was a lower quality of the medical information which also reflects in the lower validity of this diagnosis. The validity of VaD was marked as 3 (definite diagnosis) in only 31.6% of the cases, which together with No dementia (25.6%) and unspecified dementia (2.9%) were the diagnoses with the lowest proportion of cases with a validity score of 3 (Table 3). By contrast, 59.4% of the validated AD diagnoses and 50.9% of Mixed diagnoses were categorized as definite diagnosis. The quality of data retrieved for these diagnoses was categorized as 4 (good- assessed at a tertiary clinic or equal level) in 75.7% of AD diagnoses and 69.7% of the Mixed diagnoses.

During the validation process, the original register diagnosis was refined in 883 (40.0%) of 2,206 register diagnoses with dementia. The main change from an original register diagnosis was seen for unspecified dementia, where only 25% of the original NPR diagnoses remained after validation (Table 2). Unspecified dementia was primarily changed to VaD ($n = 111$, 33.3%), Mixed ($n = 82$, 24.6%), and AD ($n = 57$, 17.1%). AD was primarily re-diagnosed to Mixed dementia ($n = 227$, 78.5%), VaD to Mixed dementia ($n = 68$, 45.3%) and AD ($n = 28$, 18.7%), and Mixed dementia to AD ($n = 30$, 68.2%). For DLB, PDD, and FTD, which were all less prevalent and of higher diagnostic validity, the diagnoses were not changed to the same extent. A shift to DLB ($n = 5$, 50.0%) was seen for the PDD diagnoses that were changed. However, in cases where DLB and FTD were re-diagnosed, the distribution of new diagnoses was more varied, showing no specific pattern. Supplementary Table 1 shows all changed diagnoses.

Risk and protective factors

Baseline risk- and protective factors for the whole population, all-cause dementia, AD (Mixed included), and VaD are presented in Table 4. Comparisons were performed between the group with no dementia and the dementia groups respectively. The dementia groups were significantly older that the no dementia cases. The proportion of women was larger in AD, whereas that of men was larger in VaD in comparison with no dementia. The dementia groups had a lower level of education and a higher percentage of apolipoprotein E ε4 carriers than the no dementia group. Cholesterol levels measured as the mean S-ApoB/ApoA ratio at baseline was within normal range at baseline for the dementia groups. However, all-cause dementia and AD groups had a significantly higher mean S-ApoB/ApoA ratio than no dementia. Further, presence of hypertension, prevalent diabetes, antihypertensive medication as well as lipid lowering medication were more common in the dementia groups than in no dementia. Prevalent coronary events were more common in all-cause dementia and VaD, whereas prevalent stroke was more common only in VaD, compared with no dementia (Table 4).

DISCUSSION

In this validation study, there was a total of 2,120 dementia diagnoses after validation of 2,206 diag-

Table 3
Results on quality assessment of the medical records and validity of register diagnoses per specific diagnostic category

		Quality of data N (%)		Validity of diagnosis N (%)
Alzheimer's disease dementia	0	21 (3.4)	0	15 (2.4)
	1	20 (3.2)	1	56 (9.0)
	2	44 (7.1)	2	181 (29.1)
	3	66 (10.6)	3	369 (59.4)
	4	470 (75.7)		
Mixed dementia	0	3 (0.5)	0	3 (0.5)
	1	14 (2.4)	1	52 (8.7)
	2	46 (7.7)	2	237 (39.8)
	3	117 (19.7)	3	303 (50.9)
	4	415 (69.7)		
Vascular dementia	0	27 (5.1)	0	13 (2.4)
	1	62 (11.7)	1	117 (22.0)
	2	110 (20.7)	2	233 (43.9)
	3	86 (16.2)	3	168 (31.6)
	4	246 (46.3)		
Dementia with Lewy Bodies	0	0 (0.0)	0	0 (0.0)
	1	0 (0.0)	1	4 (4.2)
	2	1 (1.0)	2	24 (25.0)
	3	6 (6.3)	3	68 (70.8)
	4	89 (92.7)		
Parkinson disease with dementia	0	0 (0.0)	0	0 (0.0)
	1	2 (4.1)	1	4 (8.2)
	2	4 (8.2)	2	12 (24.5)
	3	8 (16.3)	3	33 (67.3)
	4	35 (71.4)		
Frontotemporal dementia	0	2 (7.7)	0	0 (0.0)
	1	0 (0.0)	1	4 (15.4)
	2	0 (0.0)	2	4 (15.4)
	3	0 (0.0)	3	18 (69.2)
	4	24 (92.3)		
Unspecified dementia	0	36 (25.7)	0	58 (41.4)
	1	30 (21.4)	1	54 (38.6)
	2	40 (28.6)	2	24 (17.1)
	3	12 (8.6)	3	34 (2.9)
	4	22 (15.7)		
No dementia	0	6 (7.0)	0	7 (8.1)
	1	10 (11.6)	1	19 (22.1)
	2	19 (22.1)	2	38 (44.2)
	3	11 (12.8)	3	22 (25.6)
	4	40 (46.5)		

Quality of data defined according to five qualitative categories; 0 = almost non-existent (only information about diagnosis); 1 = very sparse (two of the key data missing); 2 = sparse (one of the key data missing), 3 = adequate (all key data present), and 4 = good (assessed at a tertiary clinic or equal level). Validity of diagnosis defined according to four different categories; 0 = uncertain diagnosis (available data indicated different diagnoses to the same extent); 1 = possible diagnosis (some data indicated a specific diagnosis but important data was missing or the data did not fully fit into the same diagnostic criteria); 2 = probable diagnosis (a vast majority of the data indicated a specific diagnosis but some important data was missing or did not fully fit into the criteria); and 3 = definite diagnosis (all of the key data indicated the same specific diagnosis).

noses extracted from the NPR. Hence, 96% of overall dementia diagnoses have a high validity. A total of 1,323 (60.0%) original diagnoses from the NPR were unchanged during the validation process. A majority (76.9%) had been diagnosed at a Memory clinic level. For more specific diagnoses, the validity was higher than for unspecified dementia. Thus, diagnoses such as DLB and FTD had a high validity with a large proportion of the NPR diagnoses remaining unchanged after validation. Unspecified dementia had a lower diagnostic validity in the NPR, and the accessible medical information was of lower quality.

The main change from original register diagnosis was seen for unspecified dementia, where the majority of the refined cases were changed to one of the three most common dementia diagnoses: VaD, Mixed dementia, or AD. A change of diagnosis from unspecified dementia to more specific diagnoses after the

Table 4
Baseline characteristics of the MDCS presented for the total cohort, all-cause dementia, Alzheimer's dementia and Vascular dementia

Characteristic	No dementia	All-cause dementia	Alzheimer's disease dementia (Mixed included)	Vascular dementia
	N = 28 326	N = 2 120	N = 1 216	N = 531
Age mean (SD)	57.6 (7.5)	64.2 (6.0)	64.3 (5.8)	64.6 (5.9)
median (min - max)	56.9 (44.5 – 73.6)	64.6 (44.9 – 73.3)***	64.8 (44.9 – 73.3)***	64.7 (46.4 – 73.3)***
Missing data n (%)	*1 (0)*	*–*	*–*	*–*
Women	17,040 (60.2)	1,286 (60.7)	828 (68.1)***	269 (50.7)***
Education				
≤8 y	10,914 (41.2)	1,057 (53.4)***	591 (51.1)***	281 (59.4)***
9–12 y	9,319 (35.1)	646 (32.6)***	411 (35.5)***	123 (26.0)
≥13 y	6,287 (23.7)	276 (13.9)***	155 (13.4)***	69 (14.6)***
Missing data n (%)	*1,806 (6.4)*	*141 (6.7)*	*59 (4.9)*	*58 (10.9)*
Risk factors				
APOE ε4 carrier	7,854 (28.5)	1,056 (51.1)***	717 (60.4)***	206 (39.5)***
Missing data n (%)	*776 (2.7)*	*52 (2.5)*	*28 (2.3)*	*10 (1.9)*
S-ApoB/ApoA1 ratio	0.7 (0.2)	0.75 (0.22)***	0.75 (0.22)***	0.76 (0.22)
Missing data n (%)	*2,504 (8.8)*	*189 (8.9)*	*91 (7.5)*	*64 (12.1)*
Hypertension	16,127 (57.0)	1,523 (71.8)***	848 (69.7)***	431 (81.2)***
Missing data n (%)	*57 (0.2)*	*1 (0)*	*1 (0.1)*	*–*
Body mass index	25.8 (4.0)	26.2 (4.1)***	26.0 (3.9)	26.8 (4.1)***
Missing data n (%)	*53 (0.2)*	*5 (0.2)*	*1 (0)*	*3 (0.6)*
Ever smoker	16,638 (62.6)	1,107 (55.7)***	608 (52.5)***	293 (61.3)
Missing data n (%)	*1,749 (6.2)*	*133 (6.3)*	*57 (4.7)*	*53 (10.0)*
Medical history				
Prevalent diabetes	1,234 (4.4)	146 (6.9)***	76 (6.3)**	52 (9.8)***
Prevalent coronary event	538 (1.9)	62 (2.9)***	31 (2.5)	26 (4.9)***
Prevalent stroke	300 (1.1)	33 (1.6)	12 (1.0)	15 (2.8)**
Medication				
Antihypertensive medication	4,790 (16.9)	489 (23.1)***	271 (22.3)***	140 (26.4)***
Lipid lowering medication	821 (2.9)	98 (4.6)***	54 (4.4)**	27 (5.1)**

Data are presented as frequency (%) or as mean (SD). Group differences were calculated with Mann-Whitney U test and Pearson's χ^2 test with No dementia as a reference. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. BMI, body mass index; *APOE ε4*, apolipoprotein E ε4; S-ApoB/ApoA1, serum-apolipoprotein B/apolipoprotein A1. Hypertension is defined as a systolic blood pressure ≥ 140 mmHg and/or a diastolic blood pressure ≥ 90 mmHg at baseline.

validation process resulted in a larger number of diagnoses that can be used in future studies. Hence, the validation procedure led to more specific dementia diagnoses, although not all of them were classified at the highest level of diagnostic validity even after the procedure.

AD and VaD were primarily re-diagnosed to Mixed dementia, and Mixed dementia to AD, likely reflecting the difficulties to clinically distinguish between these often overlapping and symptomatically similar types of dementia. AD showed a high validity, with the majority of cases (59%) scoring 3 (definite diagnosis) reflecting the high-quality data obtained for 75% of the cases. This is likely to be a result of AD being the most prevalent dementia diagnosis, as well as a high adaptation of the diagnostic tools to identify AD-specific symptoms and signs. For VaD, the validity of the diagnosis was scored as probable in most cases, and the reported quality of data more diverse. This may be a result of the fluctuating nature

of vascular cognitive disorder symptomatology and the fact that vascular pathology often co-exists with AD pathology, especially in the older patient groups [18]. Thus, it might be difficult to clinically exclude the possibility of co-existing AD pathology, especially with lack of high-quality data. Mixed dementia had a larger proportion of diagnoses classified as definite than VaD (50.9% versus 31.6%) which is in line with the above suggestion regarding an adaption of diagnostic tools in favor of AD diagnosis. The cognitive profile indicating VaD includes reduced cognitive speed and an impaired executive functioning [15, 19] which is not always uncovered by standard cognitive screening tests such as the often-used Mini-Mental State Examination. Hence, this might result in a lower validity of VaD.

The frequency of AD in our study was 57%, if including both typical AD and Mixed diagnosis. AD is well known to be the most common pathophysiological origin of dementia comprising an estimated

frequency of 50–75% of all dementia cases which is in line with our data. Further, VaD (20%), DLB (5%), and FTD (5%) were the next most prevalent dementia diagnoses [20]. However, the commonly seen co-existence of symptomatology and pathology patterns of several dementia diagnoses make frequencies only approximations [21, 22]. To the best of our knowledge, there are no large validation studies of register-based dementia diagnosis in Sweden reporting diagnostic distribution in detail. Results from a Spanish study on validation of dementia diagnoses reported a prevalence of 58% for AD which is in line with our data [23]. However, they reported a prevalence of VaD of only 7% [23] which is low compared to 25% in our data. Differences in prevalence of dementia diagnoses between studies may be a result of various diagnostic criteria or study populations, but also whether dementia as primary or secondary diagnosis was included. It is also probable that the diagnostic registries differ between countries and hence, comparisons might not be applicable.

An underestimation of dementia diagnoses in the Swedish NPR has previously been described, partially as a result of dementia rarely being a primary cause for hospitalization, but also due to underreporting of dementia when the cause of hospitalization is another disease [24]. Another limitation of the NPR is the automatic transfer of all diagnoses without any previous quality control. However, the NPR also include policlinic specialist care such as the memory clinics which enrich the dementia diagnoses in the NPR. Nonetheless, a probable underestimation of dementia prevalence needs to be acknowledged in the present study population, since primary care diagnoses are not included unless the diagnose was registered by a hospital physician. Despite this study limitation, we believe that the register-based approach contributes to the research field by allowing large study populations and long follow-up periods. These strengths may counteract study limitations in other studies where cognitive screening and dementia assessment are part of the study protocol since such studies often are subjected to risk of attrition and loss of follow-up due to dementia, thereby suffering from health selection bias. The underestimation of dementia diagnoses in register data based on hospital discharge diagnoses has previously been described in a Finnish study, where information from public outpatient clinics as well as private clinics and hospitals was lacking [9].

Rizzuto et al. stated that misclassification between dementia subtypes is quite common in the Swedish NPR, as it also includes secondary diagnoses from hospital discharge, and suggested that this may lower the positive predictive value of AD and VaD. This may also result in a larger number of diagnoses being changed in this study, but also enabled the detection of dementia cases that would not have been found if only primary dementia diagnoses would have been included in the NPR [24]. However, there is still a substantial underdiagnosis of dementia due to a lack of public knowledge [25].

The strength of this study was its population-wide approach and a system for classification of data quality and validity of diagnoses. A limitation was a substantial variation in the accessible medical information resulting in different degrees of diagnostic quality and validity. However, unclear cases were subject to discussion with at least one experienced dementia specialist resulting in a consensus diagnosis. The possible underestimation of dementia diagnoses in the NPR as discussed previously also needs to be considered. Eventual lack of medical information was taken into consideration and, if appropriate, resulted in a lower diagnostic validity. A quality classification of diagnostic accuracy enables the study of associations between influential factors and specific dementia diagnoses with high diagnostic accuracy. This, in turn, facilitates the finding of eventual associations which may not otherwise be detected in a more heterogenous group of cases with the same dementia diagnosis. Another weakness was that the extraction of data for the study from this very extensive register was limited to individuals with a dementia diagnosis. Hence, data on cases with unrecognized dementia in the NPR was not available and the possible underestimation of dementia discussed above could not be estimated.

The risk factors/markers at baseline follow the expected pattern with an overrepresentation of women with AD and a higher proportion of apolipoprotein E $\varepsilon4$ carriers among individuals with AD, which has been reported in previous studies [26–28]. In line with previous studies, we found that hypertension [7] and diabetes mellitus [2, 29] are risk factors for vascular cognitive disorder and for AD [8]. A correlation between an increasing number of vascular risk factors and elevated levels of brain amyloid has also been described [30]. Cholesterol levels measured as S-ApoB/ApoA1 ratio was elevated in all-cause dementia and in AD, but not in VaD. Previous studies have described an association between processes related to cholesterol metabolism and AD [8] as well as elevated cholesterol levels in midlife and

VaD [7, 31]. However, covariation analyses need to be performed in order to draw conclusions regarding associations between influential factors and dementia diagnoses, which is beyond the scope of the present study.

It is of great importance to have a high diagnostic certainty when studying potential associations between different variables and dementia. As there is still no single test or examination to diagnose dementia, a thorough review of all relevant information underlying a diagnosis is therefore important in future studies using national register diagnoses of dementia. In studies of influential factors associated with any specific dementia diagnosis, only the cases with highest quality data and validity score for the diagnosis should be included in order to obtain the most accurate results.

In conclusion, whereas dementia as a general diagnosis seems to be highly valid, specific dementia diagnoses derived from registers need to be validated in order to study associations between influential factors and different specific dementia sub-type diagnoses. The results of this study facilitate further research on specific dementia diagnoses and influential factors by using data from the MDCS.

ACKNOWLEDGMENTS

FUNDING

CONFLICT OF INTEREST

The authors have no conflict of interest to report.

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/ADR-220003.

REFERENCES

[1] Vinters HV, Zarow C, Borys E, Whitman JD, Tung S, Ellis WG, Zheng L, Chui HC (2018) Review: Vascular dementia: Clinicopathologic and genetic considerations. *Neuropathol Appl Neurobiol* **44**, 247-266.

[2] Gottesman RF, Albert MS, Alonso A, Coker LH, Coresh J, Davis SM, Deal JA, McKhann GM, Mosley TH, Sharrett AR, Schneider ALC, Windham BG, Wruck LM, Knopman DS (2017) Associations between midlife vascular risk factors and 25-year incident dementia in the Atherosclerosis Risk in Communities (ARIC) Cohort. *JAMA Neurol* **74**, 1246-1254.

[3] Jorm AF, Dear KB, Burgess NM (2005) Projections of future numbers of dementia cases in Australia with and without prevention. *Aust N Z J Psychiatry* **39**, 959-963.

[4] Brookmeyer R, Gray S, Kawas C (1998) Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *Am J Public Health* **88**, 1337-1342.

[5] Norton S, Matthews FE, Barnes DE, Yaffe K, Brayne C (2014) Potential for primary prevention of Alzheimer's disease: An analysis of population-based data. *Lancet Neurol* **13**, 788-794.

[6] Ngandu T, Lehtisalo J, Solomon A, Levalahti E, Ahtiluoto S, Antikainen R, Backman L, Hanninen T, Jula A, Laatikainen T, Lindstrom J, Mangialasche F, Paajanen T, Pajala S, Peltonen M, Rauramaa R, Stigsdotter-Neely A, Strandberg T, Tuomilehto J, Soininen H, Kivipelto M (2015) A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A randomised controlled trial. *Lancet* **385**, 2255-2263.

[7] Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J (2006) Risk score for the prediction of dementia risk in 20 years among middle aged people: A longitudinal, population-based study. *Lancet Neurol* **5**, 735-741.

[8] Jones L, Holmans PA, Hamshere ML, Harold D, Moskvina V, Ivanov D, Pocklington A, Abraham R, Hollingworth P, Sims R, Gerrish A, Pahwa JS, Jones N, Stretton A, Morgan AR, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schurmann B, Heun R, Kolsch H, van den Bussche H, Heuser I, Peters O, Kornhuber J, Wiltfang J, Dichgans M, Frolich L, Hampel H, Hull M, Rujescu D, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Singleton AB, Guerreiro R, Muhleisen TW, Nothen MM, Moebus S, Jockel KH, Klopp N, Wichmann HE, Ruther E, Carrasquillo MM, Pankratz VS, Younkin SG, Hardy J, O'Donovan MC, Owen MJ, Williams J (2010) Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. *PLoS One* **5**, e13950.

[9] Solomon A, Ngandu T, Soininen H, Hallikainen MM, Kivipelto M, Laatikainen T (2014) Validity of dementia and Alzheimer's disease diagnoses in Finnish national registers. *Alzheimers Dement* **10**, 303-309.

[10] Ludvigsson JF, Almqvist C, Bonamy AK, Ljung R, Michaelsson K, Neovius M, Stephansson O, Ye W (2016) Registers of the Swedish total population and their use in medical research. *Eur J Epidemiol* **31**, 125-136.

[11] Berglund G, Elmstahl S, Janzon L, Larsson SA (1993) The Malmo Diet and Cancer Study. Design and feasibility. *J Intern Med* **233**, 45-51.

[12] Rosvall M, Ostergren PO, Hedblad B, Isacsson SO, Jan-
     zon L, Berglund G (2000) Occupational status, educational
     level, and the prevalence of carotid atherosclerosis in a gen-
     eral population sample of middle-aged Swedish men and
     women: Results from the Malmo Diet and Cancer Study.
     Am J Epidemiol **152**, 334-346.

[13] Manjer J, Carlsson S, Elmstahl S, Gullberg B, Janzon L,
     Lindstrom M, Mattisson I, Berglund G (2001) The Malmo
     Diet and Cancer Study: Representativity, cancer incidence
     and mortality in participants and non-participants. *Eur J
     Cancer Prev* **10**, 489-499.

[14] Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim
     JL, Reuterwall C, Heurgren M, Olausson PO (2011) Exter-
     nal review and validation of the Swedish national inpatient
     register. *BMC Public Health* **11**, 450.

[15] American Psychiatric Association (2013) *Diagnostic and
     statistical manual of mental disorders,* 5th ed. American
     Psychiatric Association, Washington, DC.

[16] World Health Organization (2020) Dementia Fact
     Sheet, https://www.who.int/news-room/fact-sheets/
     detail/dementia, Last updated September 2, 2021,
     Accessed on November 14, 2021.

[17] World Health Organization (2021) Hypertension Fact
     Sheet, https://www.who.int/news-room/fact-sheets/detail/
     hypertension, Last Updated August 25, 2021, Accessed on
     December 19, 2021.

[18] Langa KM, Foster NL, Larson EB (2004) Mixed demen-
     tia: Emerging concepts and therapeutic implications. *JAMA*
     292, 2901-2908.

[19] Jefferson AL, Brickman AM, Aloia MS, Paul RH (2005)
     The cognitive profile of vascular dementia. In *Vascu-
     lar Dementia, Current Clinical Neurology*, Paul RH,
     Cohen R, Ott BR, Salloway S, eds. Humana Press.
     https://doi.org/10.1385/1-59259-824-2:131

[20] Cunningham EL, McGuinness B, Herron B, Passmore AP
     (2015) Dementia. *Ulster Med J* **84**, 79-87.

[21] Knapp M, Prince M (2007) *Dementia UK: Full Report*.
     Alzheimer's Society, London.

[22] Prince M, Jackson J (2009) *World Alzheimer Report*.
     Alzheimer's Disease International, London.

[23] Andreu-Reinon ME, Gavrila D, Chirlaque MD, Colorado-
     Yohar SM, Amiano P, Ardanaz E, Navarro-Mateu F, Navarro
     C, Huerta JM (2019) Ascertainment of dementia cases in the
     Spanish European Prospective Investigation into Cancer and
     Nutrition-Murcia Cohort. *Neuroepidemiology* **52**, 63-73.

[24] Rizzuto D, Feldman AL, Karlsson IK, Dahl Aslan AK, Gatz
     M, Pedersen NL (2018) Detection of dementia cases in two
     Swedish health registers: A validation study. *J Alzheimers
     Dis* **61**, 1301-1310.

[25] Ross GW, Abbott RD, Petrovitch H, Masaki KH, Murdaugh
     C, Trockman C, Curb JD, White LR (1997) Frequency and
     characteristics of silent dementia among elderly Japanese-
     American men. The Honolulu-Asia Aging Study. *JAMA*
     277, 800-805.

[26] Riedel BC, Thompson PM, Brinton RD (2016) Age, APOE
     and sex: Triad of risk of Alzheimer's disease. *J Steroid
     Biochem Mol Biol* **160**, 134-147.

[27] Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G (2013)
     Apolipoprotein E and Alzheimer disease: Risk, mechanisms
     and therapy. *Nat Rev Neurol* **9**, 106-118.

[28] Lutz MW, Sundseth SS, Burns DK, Saunders AM, Hayden
     KM, Burke JR, Welsh-Bohmer KA, Roses AD (2016) A
     genetics-based biomarker risk algorithm for predicting risk
     of Alzheimer's disease. *Alzheimers Dement (N Y)* **2**, 30-44.

[29] Gorelick PB, Scuteri A, Black SE, Decarli C, Greenberg
     SM, Iadecola C, Launer LJ, Laurent S, Lopez OL, Nyen-
     huis D, Petersen RC, Schneider JA, Tzourio C, Arnett DK,
     Bennett DA, Chui HC, Higashida RT, Lindquist R, Nilsson
     PM, Roman GC, Sellke FW, Seshadri S; American Heart
     Association Stroke Council, Council on Epidemiology and
     Prevention, Council on Cardiovascular Nursing, Council on
     Cardiovascular Radiology and Intervention, and Council
     on Cardiovascular Surgery and Anesthesia (2011) Vascu-
     lar contributions to cognitive impairment and dementia: A
     statement for healthcare professionals from the American
     Heart Association/American Stroke Association. *Stroke* **42**,
     2672-2713.

[30] Gottesman RF, Schneider AL, Zhou Y, Coresh J, Green E,
     Gupta N, Knopman DS, Mintz A, Rahmim A, Sharrett AR,
     Wagenknecht LE, Wong DF, Mosley TH (2017) Association
     between midlife vascular risk factors and estimated brain
     amyloid deposition. *JAMA* **317**, 1443-1450.

[31] Solomon A, Kivipelto M, Wolozin B, Zhou J, Whitmer
     RA (2009) Midlife serum cholesterol and increased risk
     of Alzheimer's and vascular dementia three decades later.
     Dement Geriatr Cogn Disord **28**, 75-80.